International Journal of Research Publication and Reviews, Vol 6, Issue 6, pp 1552-1559 June 2025

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Pharma AI: An Intelligent System for Medicine Prescription

Sourabh Yadav[1], Amit Swami[2], Prathamesh Thombare[3], Vishal Junjare[4], Seema Chauhan[5]

Department of Computer Engineering, VPKBIET Baramati, Savitribai Phule Pune University, Maharashtra, India
[1]sourabh.yadav.comp.2021@vpkbiet.org, [2]amit.swami.comp.2021@vpkbiet.org, [3]prathamesh.thombare.comp.2021@vpkbiet.org,
[4]vishaljunjare@gmail.com, [5]seema.chauhan@vpkbiet.org

ABSTRACT

The intersection of artificial intelligence and healthcare has enabled unprecedented advancements in clinical decision support systems. With increasing patient loads and limited access to skilled physicians, especially in rural and under-resourced areas, there is a growing demand for intelligent solutions that aid in accurate disease diagnosis and medication prescription. This paper introduces PharmaAI, an intelligent, ML-powered medical assistant designed to predict diseases based on user-reported symptoms and recommend corresponding treatments. Leveraging machine learning models such as Random Forest and Support Vector Machine (SVM), PharmaAI classifies diseases from a curated dataset of symptom combinations and provides medicine suggestions through a structured mapping system. The project is implemented using Python and Flask, with SQLite for data persistence, and includes features such as PDF prescription generation, RESTful API support, and modular system design for deployment across web and mobile platforms. Evaluation results demonstrate an accuracy of 89.7% using Random Forest and 85.2% using SVM, with high user satisfaction in interface trials. The paper outlines the entire development lifecycle of PharmaAI, including dataset processing, model training, evaluation, system architecture, user interface design, and future prospects. The proposed system serves as a scalable and adaptable solution that bridges the diagnostic gap and enhances the reach of quality medical support.

Keywords: ***Artificial Intelligence, Machine Learning, Prescription System, Random Forest, SVM, Symptom Analysis.***

I. Introduction

A. Background

The 21st century has witnessed a tremendous rise in healthcare demands, driven by a growing population, a surge in chronic diseases, and the ongoing shortage of qualified healthcare professionals. According to the World Health Organization (WHO), nearly half of the global population lacks access to essential healthcare services. This disparity is more pronounced in developing nations and rural areas, where skilled physicians and advanced diagnostic tools are in short supply. In this context, the convergence of artificial intelligence (AI) and medicine provides a significant opportunity to transform patient care.

AI systems, especially those based on machine learning (ML), are capable of analyzing complex datasets, identifying patterns, and making predictions that can assist doctors in clinical decision-making. These systems are not intended to replace human expertise but rather to augment it, offering insights derived from vast amounts of medical data that may not be immediately apparent to a human practitioner. One of the key application areas is disease diagnosis and prescription recommendation, where AI models can evaluate symptoms and suggest likely conditions along with treatment options.

B. Problem Statement

Despite advancements in medical infrastructure in urban settings, rural and underdeveloped areas still lack adequate medical expertise. Patients are often misdiagnosed, treated with the wrong medication, or are required to travel long distances to access care. Furthermore, even in urban hospitals, physicians face an overwhelming number of patients, leading to rushed consultations and human errors. This creates a dire need for an automated system that can:

Take patient symptoms as input

Predict the most probable disease

Suggest appropriate medications

Generate formal prescriptions

Such a system must be scalable, interpretable, and easily deployable across different platforms and settings.

C. Objective

This project aims to build **PharmaAI**, an intelligent system for disease prediction and medicine prescription using machine learning. The main objectives are:

To design and train machine learning models that can accurately predict diseases based on symptoms.

To map predicted diseases to appropriate medications using rule-based logic.

To create a web-based interface using Flask where users can input symptoms and receive results in real-time.

To integrate PDF generation capabilities for producing prescription reports.

To evaluate the system for performance, usability, and accuracy.

To propose a scalable deployment model for rural health clinics, mobile apps, and hospitals.

D. Scope

PharmaAI is designed to be a modular and extensible platform. Initially, the system focuses on 40+ commonly encountered diseases and their associated symptoms, extracted from a curated dataset. Over time, the model can be retrained to accommodate new diseases, updated prescription standards, and multilingual capabilities. The system is implemented in Python and deployed via Flask as a web-based service, with backend storage using SQLite. The system is also designed with the flexibility to be expanded into mobile applications, IoT integrations, and hospital information systems (HIS).

The paper will comprehensively detail the development process, including data collection, preprocessing, model training and testing, architecture design, API generation, UI implementation, testing, evaluation, limitations, and future scope.

II. Literature Review

The development of intelligent systems in healthcare has been a growing field of research, motivated by the urgent need for accurate diagnosis, improved treatment planning, and reduced human errors in clinical decision-making. Numerous researchers have contributed to this field by integrating artificial intelligence and machine learning into medical practice. Below is a detailed review of relevant literature that has shaped the development of PharmaAI:

A. Related Work

Hijazi, S., Obeid, N., & Sabri, K. (2019) proposed a personalized medical prescription system based on logical frameworks. Their system utilized patient history and contextual information to tailor drug suggestions, demonstrating the importance of personalization in AI-driven systems [1].

Hossain, M. S., Muhammad, G., & Guizani, N. (2020) developed an Explainable AI framework that integrates mass surveillance for healthcare monitoring during pandemics like COVID-19. This work emphasized the need for transparency and ethical governance in medical AI [2].

Rajkomar, A., Dean, J., & Kohane, I. (2019) reviewed the implementation of machine learning in medical institutions, including EHR integration, predictive analytics, and triage optimization [3]. Their study serves as a foundational guideline for integrating ML in hospital systems.

Choi, E., Bahadori, M. T., Sun, J., et al. (2016) introduced RETAIN, a reverse-time attention neural network that improves the interpretability of healthcare predictions. RETAIN showcases the value of transparent AI models in clinical environments [4].

Miotto, R., Wang, F., Wang, S., et al. (2018) explored deep learning's role in patient trajectory modeling, diagnosis, and prognosis. Their work highlights scalability and complexity in deploying deep models in healthcare [5].

Esteva, A., Robicquet, A., Ramsundar, B., et al. (2019) presented an overview of how deep learning is revolutionizing disease detection and image-based diagnosis. While focusing on CNNs, their insights apply broadly to AI in medicine [6].

Beam, A. L., & Kohane, I. S. (2018) emphasized the synergy between big data and ML in healthcare and addressed common barriers such as data standardization and model bias [7].

B. Comparative Analysis

Study	Methodology	Application Area	Key Contribution	Limitation
Hijazi et al. (2019)	Logical rules + EHR	Prescription systems	Patient-specific drug recommendation	Lacks predictive modeling
Hossain et al. (2020)	Explainable AI + sensors	Public health	Real-time health surveillance framework	Limited to surveillance use cases
Rajkomar et al. (2019)	ML review (XGBoost, RF)	Hospitals/EHR	Practical ML in clinical environments	General framework, no implementation

Study	Methodology	Application Area	Key Contribution	Limitation
Choi et al. (2016)	RNN with attention	Diagnosis prediction	Interpretability in deep learning	Complexity in training
Miotto et al. (2018)	Deep Autoencoders	Prognosis	Feature extraction from large datasets	Requires large annotated datasets
Esteva et al. (2019)	CNNs	Image-based diagnosis	Deep learning success in radiology	Focused on vision-based tasks
Beam & Kohane (2018)	Data pipeline + ML	Health informatics	System-wide ML integration insights	Strategic-level paper, not technical

C. Literature Gap

Despite these advances, few systems focus specifically on end-to-end solutions that cover symptom intake, disease prediction, and drug recommendation in a lightweight, deployable framework. Most deep learning models are resource-intensive and lack explainability. Furthermore, integration into rural healthcare is rarely addressed.

PharmaAI addresses this gap by focusing on the following novel aspects:

Lightweight ML models (RF, SVM) for fast and accurate diagnosis

End-to-end pipeline from symptom to prescription

Web-based and modular design for easy deployment in rural areas

Integrated PDF prescription and user interface for usability

This review justifies the need for PharmaAI and positions it as an innovative and practical contribution to the field of intelligent healthcare systems.

III. System Requirements

The development and deployment of the PharmaAI system require a set of well-defined hardware and software components to ensure seamless functioning, scalability, and maintainability. This section outlines the minimal and recommended system specifications, development tools, and frameworks used throughout the project.

A. Hardware Requirements

Component	Minimum Requirement	Recommended Requirement
Processor	Intel i3 or equivalent	Intel i5/i7 or equivalent
RAM	4 GB	8 GB or higher
Hard Disk	250 GB HDD	500 GB SSD
Display	1024x768 resolution	Full HD resolution
Network	512 kbps internet connection	5 Mbps broadband

These specifications are applicable for both the development and client environments. For cloud deployment, any basic VM instance (AWS EC2 t2.medium or higher) suffices.

B. Development Environment

Operating System: Windows 10 / Linux Ubuntu 20.04 / macOS Monterey

IDE/Editor: Visual Studio Code (preferred), Jupyter Notebook (for model development), PyCharm

Browser Support: Chrome, Firefox, Edge

Flask-WTF: For secure form submissions (planned in future update)

This system architecture and tooling stack was selected to optimize for lightweight performance, offline capability (via SQLite), and minimal deployment barriers in resource-constrained settings such as rural clinics.

IV. System Design and Architecture

The system architecture of PharmaAI is designed to be modular, scalable, and easy to deploy in both cloud and offline environments. It consists of a multi-tier structure integrating data input, processing, inference, and output delivery.

A. High-Level Architecture

PharmaAI consists of five core modules:

User Interface Module – Facilitates input of symptoms via dropdowns and form submissions.

Backend Logic Module – Developed using Flask, it orchestrates the flow between the user interface, machine learning model, and database.

Model Inference Module – Loads the pre-trained model from model.pkl and processes encoded symptom vectors to predict diseases.

Medicine Mapping Engine – Uses the medicine_mapping.pkl file to associate the predicted disease with relevant medications.

Prescription Generator – Creates downloadable PDF prescriptions using libraries such as fpdf2.

B. Component Workflow

Input: User selects symptoms.

Encoding: Symptoms are one-hot encoded via symptom_encoder.pkl.

Prediction: Encoded vector passed to the classifier.

Output Mapping: Disease result used to look up medicine.

Prescription: A PDF report is generated.

Storage: Records are logged in the SQLite database.

V. Data Collection and Processing

A. *Dataset Source*

The dataset used in PharmaAI was compiled from publicly available healthcare repositories and simulated data based on standard symptom-disease mappings from verified medical sources such as MedlinePlus and Mayo Clinic. The final dataset contains over 5,000 entries across 40+ disease labels and over 100 symptom categories.

B. Structure of the Dataset

Each record in the dataset includes:

Patient ID (anonymized)

Binary indicators for the presence (1) or absence (0) of a list of symptoms

Target disease label

Example:

Fever	Cough	Headache	Fatigue	Disease
1	1	1	0	Influenza

VI. Model Building and Evaluation

A. Model Selection

PharmaAI uses two classification models: Random Forest (RF) and Support Vector Machine (SVM). These were chosen based on their effectiveness in handling binary feature inputs and their ability to work well on small to medium datasets.

B. Training Process

Random Forest was trained using 100 estimators with Gini impurity as the splitting criterion.

SVM used an RBF kernel and was tuned using GridSearchCV for hyperparameter optimization.

Both models were trained on 80% of the data with stratified k-fold cross-validation (k=5).

C. Evaluation Metrics

Metric	Random Forest	SVM	XGBoost
Accuracy	89.7%	85.2%	41.3%
Precision	91.3%	83.9%	43.2%
Recall	87.2%	81.5%	39.7%
F1-Score	89.2%	82.6%	41.4%

D. Bias-Variance Trade-off

Random Forest showed better generalization with lower variance.SVM occasionally underfit for rare disease categories.

E. Training vs Testing

Dataset	Random Forest	SVM	XGBoost
Training Acc.	96.2%	88.7%	49.5%
Testing Acc.	89.7%	85.2%	41.3

The above metrics validate the effectiveness of the selected models and the integrity of the training pipeline.

VII. Feature Selection and Importance

Feature selection plays a pivotal role in improving model performance and interpretability. In the context of PharmaAI, symptoms serve as binary features that signal the presence or absence of specific medical conditions. Since each symptom carries clinical significance, feature reduction was avoided. However, understanding feature importance aids in explaining model decisions and refining clinical insights.

A. Feature Importance in Random Forest

The Random Forest classifier inherently supports feature importance scoring. During training, it evaluates how each symptom contributes to the decrease in impurity (Gini index) across decision trees. The top-ranked features, based on this metric, were:

Rank	Symptom	Importance Score
1	Fever	0.165
2	Headache	0.129
3	Cough	0.114
4	Fatigue	0.098
5	Sore Throat	0.087
...	...	...

These symptoms were found to be strongly indicative across multiple disease classes.

B. Clinical Interpretation

Fever was common in over 60% of the disease records, making it a key differentiator.

Headache and Fatigue, though common, were more weighted when combined with specific clusters (e.g., headache + nausea → migraine).

Rare features like "skin rash" or "chest pain" had high class specificity and helped in disease narrowing despite lower global importance.

C. Benefits of Feature Importance Analysis

Assists in identifying redundant symptoms. Enhances model explainability

Helps clinicians understand system behavior

This feature-based insight not only builds trust with users but also contributes to continuous clinical refinement of the PharmaAI platform.

VIII. Prescription Mapping Engine

The prescription mapping engine in PharmaAI serves as the post-classification layer responsible for recommending medicines based on the predicted disease. Unlike the machine learning model used for disease prediction, the mapping system follows a rule-based logic encoded via a JSON structure.

A. Medicine Mapping Structure

After predicting a disease (e.g., "Influenza"), the system queries a dictionary object (medicine_mapping.pkl) that contains a pre-defined list of medications for each disease. Example:

```
{
  "Influenza": ["Paracetamol", "Antihistamine"],
  "Malaria": ["Artesunate", "Primaquine"]
}
```

This mapping was created based on standard treatment protocols sourced from WHO guidelines, National Health Portals, and pharmacy handbooks.

B. Rule-Based Engine Logic

Input disease name from ML model

Match it against the key in the dictionary

Return list of mapped drugs

Format and insert into prescription PDF

C. Special Handling Rules

In case of multiple predicted conditions (future update), it will prioritize based on severity ranking.

For pediatric or geriatric cases, the system can be extended to check patient metadata before selecting drug dosages (planned in v2.0).

D. Drug Database Description

The current system supports ~60 common medications linked to 40 diseases. Drug categories include:

1. Antipyretics
2. Antibacterials
3. Antimalarials
4. Antihistamines
5. Antispasmodics

E. Future Directions

Drug Interaction Checks: Avoid combinations with contraindications.

Dose Recommendations: Adjust dosage based on age, weight, allergies.

EHR Integration: Match prescription history for longitudinal tracking.

This module ensures that users receive complete and contextually accurate prescription details, streamlining the end-to-end decision support process.

IX. Mathematical Model

The PharmaAI system can be modeled using a composition of deterministic and probabilistic functions.

Let:

$S = \{s_1, s_2, ..., s_n\}$ be the input symptoms (binary vector)

$f(S)$ be the function representing the trained ML model that maps symptoms to a disease D

$g(D)$ be the medicine mapping function that outputs a list of drugs M

$h(S, D, M)$ be the function that generates a structured prescription document P

Equations:

Disease Prediction:

$D = f(S)$

Medicine Recommendation:

$M = g(D)$

Prescription Generation:

$P = h(S, D, M)$

X. Use Case Scenarios

1. Rural Clinics: Used where doctors are unavailable—staff can input symptoms and print prescriptions.

2. Emergency Ambulance Units: For pre-hospital triage and symptom recording.

3. Virtual Consultations: Embedded into telemedicine portals for fast diagnosis.

4. Urban Hospitals: As a triage assistant to reduce physician load during rush hours.

XIII. Research Contributions

End-to-End System: From symptom input to PDF prescription

Lightweight Architecture: Runs on minimal hardware

Custom Medicine Mapping: No external APIs needed

Easy Integration: Flask-based RESTful architecture

Offline Usability: Useful in non-networked areas

XIV. Results and Discussion

Best Accuracy: 89.7% (Random Forest)

AUC Score: 0.93

Error Cases: Overlap in symptom profiles for viral illnesses

User Feedback: 90%+ found interface easy to use

Limitations: Doesn't adjust dosage by age yet.

XV. Future Enhancements

Short-Term Goals:

Add multilingual support

Integrate drug interaction checker

Enable mobile version

Long-Term Goals:

Use NLP to extract symptoms from text

Integrate with wearable health trackers

Deploy as an EHR-compatible cloud API

XVII. Conclusion

PharmaAI showcases the potential of lightweight, rule-assisted machine learning systems in delivering meaningful healthcare support, particularly in environments where access to traditional medical infrastructure and expert consultation is limited or delayed. By leveraging a modular architecture combined with a Random Forest-based disease prediction model, PharmaAI achieves a high degree of diagnostic accuracy while maintaining computational efficiency suitable for low-resource settings. Its offline capabilities make it especially valuable in remote or rural areas where internet connectivity may be unreliable or absent. Furthermore, the integration of automated medicine recommendations, prescription management, and real-time order tracking within a user-friendly interface highlights its practicality for both patients and pharmacists. This system serves not only as a prototype for scalable AI applications in digital health but also as a stepping stone toward more comprehensive solutions.

XVIII. References

1. Hijazi, S., Obeid, N., & Sabri, K. (2019). On the Logical Foundation of a Personalized Medical *Prescription System*. IEEE Access, 7, 2020–2030.
2. Hossain, M. S., Muhammad, G., & Guizani, N. (2020). *Explainable AI and Mass Surveillance System-Based Healthcare Framework*. IEEE Network, 34(4), 126–132. https://doi.org/10.1109/MNET.011.2000125
3. Rajkomar, A., Dean, J., & Kohane, I. (2019). *Machine Learning in Medicine*. NEJM, 380, 1347–1358.
4. Choi, E., et al. (2016). *RETAIN: An Interpretable Predictive Model for Healthcare*. NeurIPS. https://arxiv.org/abs/1608.05745
5. Miotto, R., et al. (2018). *Deep Learning for Healthcare: Review*. Briefings in Bioinformatics, 19(6), 1236–1246.
6. Esteva, A., et al. (2019). *A Guide to Deep Learning in Healthcare*. Nature Medicine, 25, 24–29.
7. Topol, E. (2019). *Deep Medicine: How AI Can Make Healthcare Human Again*. Basic Books.
8. Beam, A. L., & Kohane, I. S. (2018). *Big Data and Machine Learning in Health Care*. JAMA, 319(13), 1317–1318.

XIX. Appendices

- Appendix A: Model Confusion Matrices and ROC Graphs
- Appendix B: Screenshot Walkthrough of Web UI
- Appendix C: Database Schema and SQL Dump (structure only)
- Appendix D: Sample Prescriptions (de-identified)
- Appendix E: API Documentation Snapshots
- Appendix F: Legal Disclaimer Template